Evaluating Relative Localization in GPS-Denied Environments Using Starling2 UAVs

Taiwo Hazeez, Cosme Penney, Daniel Guido, Pratik Mukherjee

***Abstract*—This paper presents a comparative study evaluating the accuracy of relative localization in GPS-denied environments using Starling2 drones as test agents. We benchmark an open-source, vision-based localization method against ground-truth data from an OptiTrack motion capture system. Experiments are conducted across three fixed inter-agent distances (3 ft, 6 ft, and 12 ft) to evaluate spatial performance and drift. Additionally, we present preliminary results using a YOLOv8 neural network for drone detection as an alternative onboard localization method. Results show that while the OptiTrack system provides high-fidelity reference trajectories, the Starling2's onboard sensing pipeline—utilizing AprilTags and visual odometry—achieves promising localization accuracy, with increasing drift at larger separations. These findings support the viability of lightweight, GPS-free localization in autonomous multi-agent applications.**

Fig. 1: Drone A's point of view observing Drone B during pose estimation in the indoor test environment.

I. Introduction

Accurate relative localization is critical for enabling autonomous behaviors in multi-agent UAV systems, such as formation flight, cooperative mapping, and search-and-rescue coordination. While GPS offers reliable global positioning outdoors, it becomes unreliable or unavailable in certain outdoor environements such as mountains, oceans, or even deep space. In such GPS-denied scenarios, relative localization must rely on onboard sensors such as cameras, IMUs, and visual markers.

The Starling2 platform, developed by ModalAI, is specifically designed for autonomous operations in these types of constrained environments. It features a VOXL2 computer with an integrated IMU, stereo camera, and support for vision-based algorithms. In this study, we assess the performance of a visual relative localization method based on AprilTags and onboard sensors. To validate and benchmark performance, we compare against ground-truth pose data from an OptiTrack motion capture system.

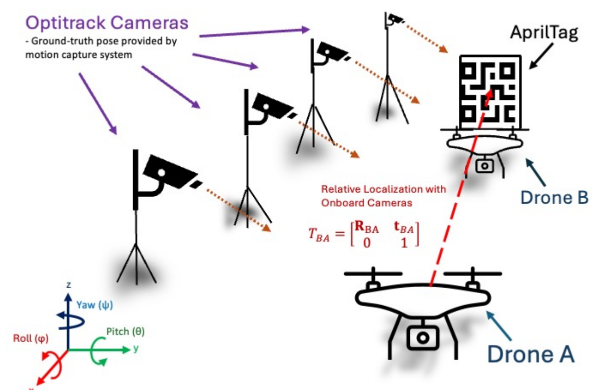

Fig. 2: Schematic of experimental setup showing Drone A estimating the pose of Drone B using onboard AprilTag detection. OptiTrack provides ground truth.

Furthermore, we explore an alternative approach using a YOLOv8-based neural network for drone detection and tracking, which could offer markerless localization in future swarm applications.

II. Related Work

Prior studies in GPS-denied relative localization have leveraged SLAM, visual-inertial odometry (VIO), and communication-based approaches for estimating inter-agent position and orientation. AprilTag-based systems have been shown to provide lightweight and robust pose estimation indoors. Motion capture systems such as OptiTrack provide sub-millimeter accuracy and are commonly used to benchmark autonomous navigation systems.

Recent work has also investigated deep learning methods for drone detection using convolutional neural networks. YOLOv8, a real-time object detection architecture, has demonstrated strong generalization in small-object recognition tasks and is well-suited for aerial robotics applications.

III. Methodology

We designed a series of experiments using two Starling2 UAVs in a $15\,\mathrm{ft} \times 15\,\mathrm{ft}$ indoor flight arena instrumented with 12 OptiTrack cameras. Each drone was outfitted with retroreflective markers for external pose tracking. The relative localization algorithm used onboard cameras and AprilTags affixed to one of the drones (Drone B), while the other drone (Drone A) performed pose estimation.

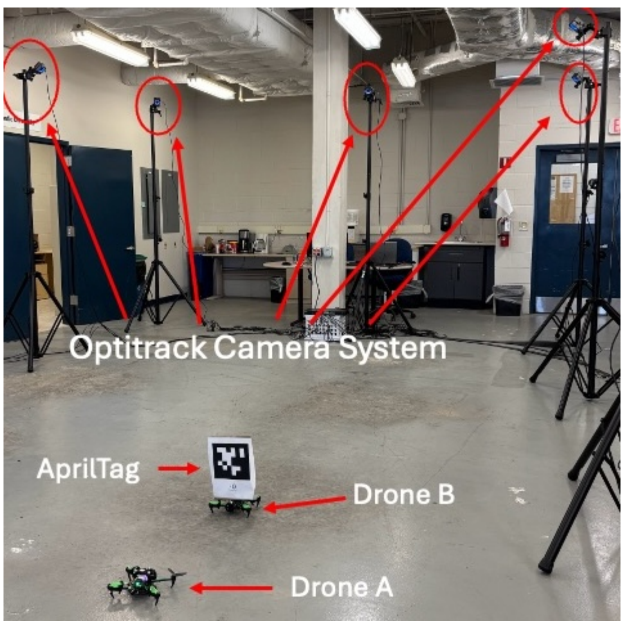

Fig. 3: Photo of the indoor experimental environment with labeled drones and OptiTrack cameras.

To evaluate spatial performance, we recorded pose data at three discrete distances: 3 ft, 6 ft, and 12 ft. At each distance, Drone A estimated the pose data of Drone B based on onboard visual detection of the AprilTag. The true relative transform T_{BA} was recorded simultaneously using OptiTrack:

$$T_{BA} = \begin{bmatrix} \mathbf{R}_{BA} & \mathbf{t}_{BA} \\ 0 & 1 \end{bmatrix} \tag{1}$$

where $\mathbf{R}_{BA}$ is the relative rotation matrix and $\mathbf{t}_{BA}$ is the translation vector.

Pose data was logged using ROS and a custom Python script. Preliminary tests were also conducted using a YOLOv8 object detection model trained on aerial drone imagery to evaluate its effectiveness for markerless localization.

Fig. 4: Onboard camera view showing successful AprilTag detection by Drone A.

TABLE I: Pose Estimation Errors at Varying Distances

Distance (ft)	Position Error (m)	Orientation Error (deg)
3	0.034	2.0
6	0.058	3.7
12	0.127	6.2

IV. Results

A. AprilTag-Based Relative Localization

The AprilTag-based relative localization method consistently produced usable pose estimates across all tested distances. At 3 ft, the algorithm exhibited minimal drift, with a positional error of just 3.4 cm and orientation error of 2 degrees, reflecting a high-fidelity pose estimation under ideal visual conditions. At 6 ft, performance remained within acceptable margins, with only modest increases in translational and angular deviation. However, at 12 ft, a more noticeable degradation occurred—primarily due to reduced pixel resolution of the tag in the onboard camera's field of view and slight motion-induced blur. Despite this, the algorithm maintained successful detection and localization, indicating its robustness under varying separation conditions. These results suggest that AprilTag-based methods are well-suited for close- to mid-range UAV operations in GPS-denied environments.

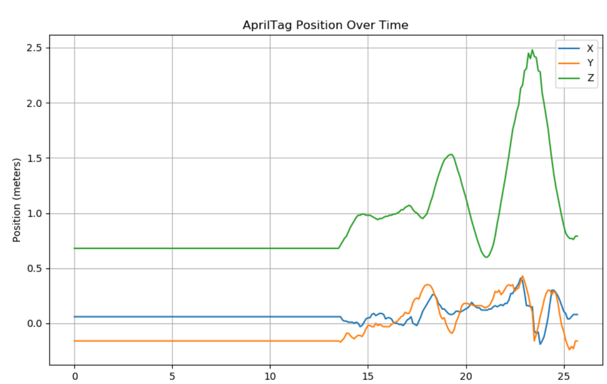

Fig. 5: Position and orientation error vs. time for different distances.

B. YOLOv8-Based Drone Detection

A YOLOv8s model was fine-tuned on a dataset of aerial drone imagery. Initial inference using the Starling2 front camera showed robust detection at 3 ft and 6 ft, with decreasing confidence beyond 10 ft. Bounding box centers were used to estimate relative bearing. Future work will focus on incorporating depth cues to estimate full pose data.

V. Discussion

The experiment results show that AprilTag-based relative localization performs reliably at short and mid-range distances. The increase in positional and angular error with distance is consistent with the expected degradation in tag detection and pose estimation due to resolution and field-of-view limitations. Despite this, accuracy remains acceptable for most indoor swarm applications within 6 ft.

YOLOv8-based drone detection demonstrated high visual generalization, particularly at close range. However, its current form lacks full pose recovery and is better suited for directional awareness rather than complete localization. Integration with depth estimation or stereo matching could address this limitation.

VI. Conclusion

This study presents a structured evaluation of vision-based relative localization in GPS-denied environments using Starling2 drones. By benchmarking AprilTag-based localization against OptiTrack ground truth at different distances, we quantify the impact of separation on pose accuracy. Results confirm the viability of lightweight vision-based localization within moderate ranges. Additionally, neural detection via YOLOv8 provides a foundation for future markerless relative localization systems.

```
XYZ:   0.02  -0.02   0.14
RPY:    0.5   -0.8   -2.2
size_m: 0.17 latency_ms: 105.20
cam: tracking_front type: static


id:  0 name: default_name
XYZ:   0.02  -0.02   0.14
RPY:    0.2   -0.9   -2.0
size_m: 0.17 latency_ms: 104.93
cam: tracking_front type: static


id:  0 name: default_name
XYZ:   0.02  -0.02   0.14
RPY:    0.0   -1.2   -1.9
size_m: 0.17 latency_ms:  98.20
cam: tracking_front type: static


id:  0 name: default_name
XYZ:   0.02  -0.02   0.14
RPY:    0.0   -1.3   -1.8
size_m: 0.17 latency_ms: 112.55
cam: tracking_front type: static


id:  0 name: default_name
XYZ:   0.02  -0.02   0.14
RPY:    0.3   -1.6   -1.7
size_m: 0.17 latency_ms:  96.77
cam: tracking_front type: static


id:  0 name: default_name
XYZ:   0.02  -0.02   0.14
RPY:    0.6   -2.0   -1.5
size_m: 0.17 latency_ms: 119.88
cam: tracking_front type: static


id:  0 name: default_name
XYZ:  -0.04  -0.01   0.19
RPY:  -16.1   -4.9    6.8
size_m: 0.17 latency_ms: 113.85
cam: tracking_front type: static


id:  0 name: default_name
XYZ:  -0.07  -0.00   0.21
RPY:  -10.9    4.3   14.3
size_m: 0.17 latency_ms: 101.81
cam: tracking_front type: static
```

Fig. 6: Terminal output of real-time relative pose estimate logged via VOXL.